S. GUCKER RE-RUN

OIPE

#34

RAW SEQUENCE LISTING DATE: 11/25/2003
PATENT APPLICATION: **US/08/879,469** TIME: 07:40:36

Input Set : **N:\Crf3\RULE60\08879469.RAW.txt**
Output Set: **N:\CRF4\11252003\H879469.raw**

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Murgita, Robert A.

(ii) TITLE OF INVENTION: RECOMBINANT HUMAN ALPHA-FETOPROTEIN AS A CELL PROLIFERATIVE AGENT

(iii) NUMBER OF SEQUENCES: 16

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 225 Franklin Street, Suite 3100
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110-2804

ENTERED

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
C--> **(A) APPLICATION NUMBER: US/08/879,469**
C--> **(B) FILING DATE: 20-Jun-1997**
W--> **(C) CLASSIFICATION: 435**

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US/08/377,316
(B) FILING DATE: 24-JAN-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Clark, Paul T.
(B) REGISTRATION NUMBER: 30,162
(C) REFERENCE/DOCKET NUMBER: 06727/006001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 542-5070
(B) TELEFAX: (617) 542-8906
(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2022 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

W--> **(ii) MOLECULE TYPE: DNA**

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATATTGTGCT TCCACCACTG CCAATAACAA AATAACTAGC AACCATGAAG TGGGTGGAAT 60
CAATTTTTTT AATTTTCCTA CTAAATTTTA CTGAATCCAG AACACTGCAT AGAAATGAAT 120

RAW SEQUENCE LISTING DATE: 11/25/2003
PATENT APPLICATION: **US/08/879,469** TIME: 07:40:36

Input Set : **N:\Crf3\RULE60\08879469.RAW.txt**
Output Set: **N:\CRF4\11252003\H879469.raw**

```
ATGGAATAGC TTCCATATTG GATTCTTACC AATGTACTGC AGAGATAAGT TTAGCTGACC     180
TGGCTACCAT ATTTTTTGCC CAGTTTGTTC AAGAAGCCAC TTACAAGGAA GTAAGCAAAA     240
TGGTGAAAGA TGCATTGACT GCAATTGAGA AACCCACTGG AGATGAACAG TCTTCAGGGT     300
GTTTAGAAAA CCAGCTACCT GCCTTTCTGG AAGAACTTTG CCATGAGAAA GAAATTTTGG     360
AGAAGTACGG ACATTCAGAC TGCTGCAGCC AAAGTGAAGA GGGAAGACAT AACTGTTTTC     420
TTGCACACAA AAAGCCCACT GCAGCATGGA TCCCACTTTT CCAAGTTCCA GAACCTGTCA     480
CAAGCTGTGA AGCATATGAA GAAGACAGGG AGACATTCAT GAACAAATTC ATTTATGAGA     540
TAGCAAGAAG GCATCCCTTC CTGTATGCAC CTACAATTCT TCTTTCGGCT GCTGGGTATG     600
AGAAAATAAT TCCATCTTGC TGCAAAGCTG AAAATGCAGT TGAATGCTTC CAAACAAAGG     660
CAGCAACAGT TACAAAAGAA TTAAGAGAAA GCAGCTTGTT AAATCAACAT GCATGTCCAG     720
TAATGAAAAA TTTTGGGACC CGAACTTTCC AAGCCATAAC TGTTACTAAA CTGAGTCAGA     780
AGTTTACCAA AGTTAATTTT ACTGAAATCC AGAAACTAGT CCTGGATGTG GCCCATGTAC     840
ATGAGCACTG TTGCAGAGCA GATGTGCTGG ATTGTCTGCA GGATGGGGAA AAAATCATGT     900
CCTACATATG TTCTCAACAA GACACTCTGT CAAACAAAAT AACAGAATGC TGCAAACTGA     960
CCACGCTGGA ACGTGGTCAA TGTATAATTC ATGCAGAAAA TGATGAAAAA CCTGAAGGTC    1020
TATCTCCAAA TCTAAACAGG TTTTTAGGAG ATAGAGATTT TAACCAATTT TCTTCAGGGG    1080
AAAAAAATAT CTTCTTGGCA AGTTTTGTTC ATGAATATTC AAGAAGACAT CCTCAGCTTG    1140
CTGTCTCAGT AATTCTAAGA GTTGCTAAAG GATACCAGGA GTTATTGGAG AAGTGTTTCC    1200
AGACTGAAAA CCCTCTTGAA TGCCAAGATA AAGGAGAAGA AGAATTACAG AAATACATCC    1260
AGGAGAGCCA AGCATTGGCA AAGCGAAGCT GCGGCCTCTT CCAGAAACTA GGAGAATATT    1320
ACTTACAAAA TGAGTTTCTC GTTGCTTACA CAAAGAAAGC CCCCCAGCTG ACCTCGTCGG    1380
AGCTGATGGC CATCACCAGA AAAATGGCAG CCACAGCAGC CACTTGTTGC CAACTCAGTG    1440
AGGACAAACT ATTGGCCTGT GGCGAGGGAG CGGCTGACAT TATTATCGGA CACTTATGTA    1500
TCAGACATGA AATGACTCCA GTAAACCCTG GTGTTGGCCA GTGCTGCACT TCTTCATATG    1560
CCAACAGGAG GCCATGCTTC AGCAGCTTGG TGGTGGATGA AACATATGTC CCTCCTGCAT    1620
TCTCTGATGA CAAGTTCATT TTCCATAAGG ATCTGTGCCA AGCTCAGGGT GTAGCGCTGC    1680
AAAGGATGAA GCAAGAGTTT CTCATTAACC TTGTGAAGCA AAAGCCACAA ATAACAGAGG    1740
AACAACTTGA GGCTCTCATT GCAGATTTCT CAGGCCTGTT GGAGAAATGC TGCCAAGGCC    1800
AGGAACAGGA AGTCTGCTTT GCTGAAGAGG GACAAAAACT GATTTCAAAA ACTGGTGCTG    1860
CTTTGGGAGT TTAAATTACT TCAGGGGAAG AGAAGACAAA ACGAGTCTTT CATTCGGTGT    1920
GAACTTTTCT CTTTAATTTT AACTGATTTA ACACTTTTTG TGAATTAATG ATAAAGACTT    1980
TTATGTGAGA TTTCCTTATC ACAGAAATAA AATATCTCCA AA                       2022
(2) INFORMATION FOR SEQ ID NO: 2:
     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 590 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: Not Relevant
          (D) TOPOLOGY: linear
    (ii) MOLECULE TYPE: protein
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
     Thr Leu His Arg Asn Glu Tyr Gly Ile Ala Ser Ile Leu Asp Ser Tyr
     1               5                  10                  15
     Gln Cys Thr Ala Glu Ile Ser Leu Ala Asp Leu Ala Thr Ile Phe Phe
                 20                  25                  30
     Ala Gln Phe Val Gln Glu Ala Thr Tyr Lys Glu Val Ser Lys Met Val
             35                  40                  45
     Lys Asp Ala Leu Thr Ala Ile Glu Lys Pro Thr Gly Asp Glu Gln Ser
         50                  55                  60
     Ser Gly Cys Leu Glu Asn Gln Leu Pro Ala Phe Leu Glu Glu Leu Cys
```

RAW SEQUENCE LISTING DATE: 11/25/2003
PATENT APPLICATION: **US/08/879,469** TIME: 07:40:36

Input Set : **N:\Crf3\RULE60\08879469.RAW.txt**
Output Set: **N:\CRF4\11252003\H879469.raw**

```
158      65                  70                  75                  80
160      His Glu Lys Glu Ile Leu Glu Lys Tyr Gly His Ser Asp Cys Cys Ser
161                      85                  90                  95
163      Gln Ser Glu Glu Gly Arg His Asn Cys Phe Leu Ala His Lys Lys Pro
164                  100                 105                 110
166      Thr Ala Ala Trp Ile Pro Leu Phe Gln Val Pro Glu Pro Val Thr Ser
167              115                 120                 125
169      Cys Glu Ala Tyr Glu Glu Asp Arg Glu Thr Phe Met Asn Lys Phe Ile
170          130                 135                 140
172      Tyr Glu Ile Ala Arg Arg His Pro Phe Leu Tyr Ala Pro Thr Ile Leu
173      145                 150                 155                 160
175      Leu Ser Ala Ala Gly Tyr Glu Lys Ile Ile Pro Ser Cys Cys Lys Ala
176                      165                 170                 175
178      Glu Asn Ala Val Glu Cys Phe Gln Thr Lys Ala Ala Thr Val Thr Lys
179                  180                 185                 190
181      Glu Leu Arg Glu Ser Ser Leu Leu Asn Gln His Ala Cys Pro Val Met
182              195                 200                 205
184      Lys Asn Phe Gly Thr Arg Thr Phe Gln Ala Ile Thr Val Thr Lys Leu
185          210                 215                 220
187      Ser Gln Lys Phe Thr Lys Val Asn Phe Thr Glu Ile Gln Lys Leu Val
188      225                 230                 235                 240
190      Leu Asp Val Ala His Val His Glu His Cys Cys Arg Ala Asp Val Leu
191                      245                 250                 255
193      Asp Cys Leu Gln Asp Gly Glu Lys Ile Met Ser Tyr Ile Cys Ser Gln
194                  260                 265                 270
196      Gln Asp Thr Leu Ser Asn Lys Ile Thr Glu Cys Cys Lys Leu Thr Thr
197              275                 280                 285
199      Leu Glu Arg Gly Gln Cys Ile Ile His Ala Glu Asn Asp Glu Lys Pro
200          290                 295                 300
202      Glu Gly Leu Ser Pro Asn Leu Asn Arg Phe Leu Gly Asp Arg Asp Phe
203      305                 310                 315                 320
205      Asn Gln Phe Ser Ser Gly Glu Lys Asn Ile Phe Leu Ala Ser Phe Val
206                      325                 330                 335
208      His Glu Tyr Ser Arg Arg His Pro Gln Leu Ala Val Ser Val Ile Leu
209                  340                 345                 350
211      Arg Val Ala Lys Gly Tyr Gln Glu Leu Leu Glu Lys Cys Phe Gln Thr
212              355                 360                 365
214      Glu Asn Pro Leu Glu Cys Gln Asp Lys Gly Glu Glu Glu Leu Gln Lys
215          370                 375                 380
217      Tyr Ile Gln Glu Ser Gln Ala Leu Ala Lys Arg Ser Cys Gly Leu Phe
218      385                 390                 395                 400
220      Gln Lys Leu Gly Glu Tyr Tyr Leu Gln Asn Glu Phe Leu Val Ala Tyr
221                      405                 410                 415
223      Thr Lys Lys Ala Pro Gln Leu Thr Ser Ser Glu Leu Met Ala Ile Thr
224                  420                 425                 430
226      Arg Lys Met Ala Ala Thr Ala Ala Thr Cys Cys Gln Leu Ser Glu Asp
227              435                 440                 445
229      Lys Leu Leu Ala Cys Gly Glu Gly Ala Ala Asp Ile Ile Ile Gly His
230          450                 455                 460
```

RAW SEQUENCE LISTING DATE: 11/25/2003
PATENT APPLICATION: US/08/879,469 TIME: 07:40:36

Input Set : N:\Crf3\RULE60\08879469.RAW.txt
Output Set: N:\CRF4\11252003\H879469.raw

```
232     Leu Cys Ile Arg His Glu Met Thr Pro Val Asn Pro Gly Val Gly Gln
233     465                 470                 475                 480
235     Cys Cys Thr Ser Ser Tyr Ala Asn Arg Arg Pro Cys Phe Ser Ser Leu
236                     485                 490                 495
238     Val Val Asp Glu Thr Tyr Val Pro Pro Ala Phe Ser Asp Asp Lys Phe
239                 500                 505                 510
241     Ile Phe His Lys Asp Leu Cys Gln Ala Gln Gly Val Ala Leu Gln Arg
242             515                 520                 525
244     Met Lys Gln Glu Phe Leu Ile Asn Leu Val Lys Gln Lys Pro Gln Ile
245         530                 535                 540
247     Thr Glu Glu Gln Leu Glu Ala Leu Ile Ala Asp Phe Ser Gly Leu Leu
248     545                 550                 555                 560
250     Glu Lys Cys Cys Gln Gly Gln Glu Gln Glu Val Cys Phe Ala Glu Glu
251                     565                 570                 575
253     Gly Gln Lys Leu Ile Ser Lys Thr Gly Ala Ala Leu Gly Val
254                 580                 585                 590
256 (2) INFORMATION FOR SEQ ID NO: 3:
258     (i) SEQUENCE CHARACTERISTICS:
259          (A) LENGTH: 197 amino acids
260          (B) TYPE: amino acid
261          (C) STRANDEDNESS: Not Relevant
262          (D) TOPOLOGY: linear
264     (ii) MOLECULE TYPE: protein
269     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
271     Thr Leu His Arg Asn Glu Tyr Gly Ile Ala Ser Ile Leu Asp Ser Tyr
272     1                   5                   10                  15
274     Gln Cys Thr Ala Glu Ile Ser Leu Ala Asp Leu Ala Thr Ile Phe Phe
275                 20                  25                  30
277     Ala Gln Phe Val Gln Glu Ala Thr Tyr Lys Glu Val Ser Lys Met Val
278             35                  40                  45
280     Lys Asp Ala Leu Thr Ala Ile Glu Lys Pro Thr Gly Asp Glu Gln Ser
281         50                  55                  60
283     Ser Gly Cys Leu Glu Asn Gln Leu Pro Ala Phe Leu Glu Glu Leu Cys
284     65                  70                  75                  80
286     His Glu Lys Glu Ile Leu Glu Lys Tyr Gly His Ser Asp Cys Cys Ser
287                     85                  90                  95
289     Gln Ser Glu Glu Gly Arg His Asn Cys Phe Leu Ala His Lys Lys Pro
290                 100                 105                 110
292     Thr Ala Ala Trp Ile Pro Leu Phe Gln Val Pro Glu Pro Val Thr Ser
293             115                 120                 125
295     Cys Glu Ala Tyr Glu Glu Asp Arg Glu Thr Phe Met Asn Lys Phe Ile
296         130                 135                 140
298     Tyr Glu Ile Ala Arg Arg His Pro Phe Leu Tyr Ala Pro Thr Ile Leu
299     145                 150                 155                 160
301     Leu Ser Ala Ala Gly Tyr Glu Lys Ile Ile Pro Ser Cys Cys Lys Ala
302                     165                 170                 175
304     Glu Asn Ala Val Glu Cys Phe Gln Thr Lys Ala Ala Thr Val Thr Lys
305                 180                 185                 190
307     Glu Leu Arg Glu Ser
```

RAW SEQUENCE LISTING DATE: 11/25/2003
PATENT APPLICATION: **US/08/879,469** TIME: 07:40:36

Input Set : **N:\Crf3\RULE60\08879469.RAW.txt**
Output Set: **N:\CRF4\11252003\H879469.raw**

195

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: Not Relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Ser Leu Leu Asn Gln His Ala Cys Pro Val Met Lys Asn Phe Gly Thr
1               5                  10                  15
Arg Thr Phe Gln Ala Ile Thr Val Thr Lys Leu Ser Gln Lys Phe Thr
            20                  25                  30
Lys Val Asn Phe Thr Glu Ile Gln Lys Leu Val Leu Asp Val Ala His
        35                  40                  45
Val His Glu His Cys Cys Arg Ala Asp Val Leu Asp Cys Leu Gln Asp
    50                  55                  60
Gly Glu Lys Ile Met Ser Tyr Ile Cys Ser Gln Gln Asp Thr Leu Ser
65                  70                  75                  80
Asn Lys Ile Thr Glu Cys Cys Lys Leu Thr Thr Leu Glu Arg Gly Gln
                85                  90                  95
Cys Ile Ile His Ala Glu Asn Asp Glu Lys Pro Glu Gly Leu Ser Pro
            100                 105                 110
Asn Leu Asn Arg Phe Leu Gly Asp Arg Asp Phe Asn Gln Phe Ser Ser
        115                 120                 125
Gly Glu Lys Asn Ile Phe Leu Ala Ser Phe Val His Glu Tyr Ser Arg
    130                 135                 140
Arg His Pro Gln Leu Ala Val Ser Val Ile Leu Arg Val Ala Lys Gly
145                 150                 155                 160
Tyr Gln Glu Leu Leu Glu Lys Cys Phe Gln Thr Glu Asn Pro Leu Glu
                165                 170                 175
Cys Gln Asp Lys Gly Glu Glu Glu Leu Gln Lys Tyr Ile Gln Glu Ser
            180                 185                 190
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 201 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: Not Relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Gln Ala Leu Ala Lys Arg Ser Cys Gly Leu Phe Gln Lys Leu Gly Glu
1               5                  10                  15
Tyr Tyr Leu Gln Asn Glu Phe Leu Val Ala Tyr Thr Lys Lys Ala Pro
            20                  25                  30
Gln Leu Thr Ser Ser Glu Leu Met Ala Ile Thr Arg Lys Met Ala Ala
        35                  40                  45
Thr Ala Ala Thr Cys Cys Gln Leu Ser Glu Asp Lys Leu Leu Ala Cys
    50                  55                  60
```

VERIFICATION SUMMARY DATE: 11/25/2003
PATENT APPLICATION: **US/08/879,469** TIME: 07:40:37

Input Set : **N:\Crf3\RULE60\08879469.RAW.txt**
Output Set: **N:\CRF4\11252003\H879469.raw**

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:33 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:55 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:682 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:699 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:716 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:733 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12
L:750 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=13
L:767 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14
L:784 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15